

МАСШТАБНО ИНВАРИАНТНАЯ КОЛИЧЕСТВЕННАЯ ОЦЕНКА ШРИФТА, ОСНОВАННАЯ НА ИДЕЯХ ФРАКТАЛЬНОЙ ГЕОМЕТРИИ

¹Тарасов Д.А., ¹Емелхан А.Б., ¹Сергеев А.П., ¹Тягунов А.Г.

¹ФГАУ ВПО «Уральский Федеральный Университет имени первого Президента России Б.Н. Ельцина», Россия, Екатеринбург, Мира-32, 620002, E-mail: datarasov@yandex.ru

Для оценки пространственных особенностей шрифтовых рисунков предложено использовать масштабно инвариантный показатель, основанный на идеях фрактальной геометрии. Шрифт рассматривается как связанное геометрическое множество векторных графических объектов и к нему применяется определение нормализованной компактности. Для вычисления периметра и площади векторного объекта использован VBA макрос Curveinfo векторного графического пакета CorelDraw. Был разработан специальный макрос, который автоматизирует процесс разбиения произвольного количества текстовых объектов на набор векторных объектов представляющих из себя абрисы букв и внутрибуквенные просветы. Выборка: 21 шрифт, размеры 12 и 18 пт; из них 9 шрифтов без засечек, 11 шрифтов с засечками и один скриптовый шрифт. Результаты и измерений и расчетов показали масштабную инвариантность нормализованной компактности и возможность использования этого показателя как численной характеристики и основания для классификации шрифтов.

Ключевые слова: текст, шрифт, скорость чтения, показатель, фрактал,.

A SUBSTANTIAL SCALE INVARIANT QUANTITATIVE EVALUATION OF A FONT BASED ON IDEAS OF FRACTAL GEOMETRY

¹Tarasov D.A., ¹Yemelkhan A.B., ¹Sergeev A.P., ¹Tyagunov A.G.

¹Ural Federal University, Russia, Yekaterinburg, Mira-32, 620002, E-mail: datarasov@yandex.ru

For an assessment of spatial features of font drawings it is offered to use on a substantial scale invariant indicator based on ideas of fractal geometry. The font is considered as a coherent geometrical set of vector graphic objects and determination of the normalized compactness is applied to it. For calculation of perimeter and the area of vector object VBA a macro of Curelinfo of a vector graphic CorelDraw package is used. The special macro which automates process of splitting any amount of text objects into a set of vector objects outlines of the letters and inside the alphabetic gleams which are selection was developed: 21 fonts, sizes 12 and 18 point; from them 9 sans-serif fonts, 11 serif types and one scripting font. Results and measurements and calculations showed large-scale invariancy of the normalized compactness and possibility of use of this indicator as numerical characteristic and basis for classification of fonts.

Key words: text, index, reading speed font, fractal,

Введение

Исследования в области разборчивости и читабельности текстовой информации ведутся на протяжении более ста лет. Они необходимы для разработки текстовых материалов, предназначенных для читателей с формирующимся навыком чтения. Значительное место в них занимают исследования шрифтов. Многие зарубежные и отечественные авторы исследовали разборчивость, различимость, удобство чтения различных шрифтов, влияние засечек, влияние рисунка и пространственных характеристик шрифта на понимание и запоминание содержания текстовых информационных, и некоторые другие факторы. Полученные результаты противоречивы и до сих пор нет единого мнения о том, какие особенности шрифтов и каким образом влияют на процесс чтения. Во многом это

связано с отсутствием объективного показателя, описывающего рисунок шрифта и позволяющего сравнивать различные шрифты между собой.

Цель исследования

Предложить масштабно инвариантный показатель, численно характеризующий пространственные характеристики шрифта. Он предназначен для количественной оценки разборчивости шрифтов.

Текущее состояние вопроса

Различия в удобочитаемости гарнитур исследовались в [1]. Продемонстрировано незначительное превосходство ряда гарнитур, связанное с их рисунком и начертанием. При этом субъективные оценки респондентов были часто противоположными. Аналогичные результаты были получены в [2] при разработке новых шрифтов. Работы показали наличие субъективных предпочтений читателей, а также объективной разности в удобочитаемости гарнитур различных начертаний. В работе [3] исследования удобочитаемости шрифтов осуществлялись сравнением скорости чтения отрывков текстовых информационных сообщений на русском языке, по 2000 знаков, набранных некоторыми распространенными шрифтами, кеглем 12 пт. Проведена кластеризация шрифтов на основе количества отмеченных респондентами меток и произведено ранжирование шрифтов по удобочитаемости. Продемонстрирована более высокая скорость чтения на шрифтах с засечками. Однако каких-либо явных характеристик шрифтов, влияющих на удобочитаемость, выявлено не было. В [4] дан обзор ситуации с современным шрифтовым оформлением учебников для средней школы. Рассмотрено противоречие норм действующего технического регулирования и современного состояния шрифтового оформления. Выявлен дефицит гарнитур, предназначенных для набора учебной литературы. Показано, что санитарные нормы определяют требования лишь к высоте прописных символов, что ни в коей мере не отражает реального положения с удобочитаемостью шрифта, поскольку основной вклад в нее делает высота строчных символов. Сделано заключение о том, что в России практически отсутствуют современные исследования удобочитаемости.

В фундаментальной работе «*Legibility of print*» Майлза Тинкера [5] показано, что при чтении строчные буквы более важны, чем заглавные, а разборчивость некоторых букв может быть улучшена за счет использования засечек, насыщенности штрихов, разграничения отличительных характеристик, упрощения контура, пустого пространства в очке литеры, ширины буквы.

Известный британский исследователь чтения и восприятия Э. Дж. Сэнфорд подтвердил, что шрифты различаются по своей разборчивости [6]. Читабельность зависит от начертания символов гарнитуры, определяющую разборчивость слов и букв. Таким образом, большая разборчивость приводит к большей способности различать и, следовательно, в большем умении читать текстовых информации [7]. В [8], был проведен анализ публикаций, касающихся роли букв в восприятии текстовых информации в целом и показано, что начертание отдельных символов и их идентификация мозгом, их пространственные характеристики и взаимодействие между собой, равно как и шрифты, и гарнитуры в целом, являются существенным фактором, влияющим на чтение, требующим дальнейших исследований.

Подход и инструментарий

Оценка визуальных характеристик шрифтов представляет определенные трудности, связанные с различием в понимании того, что представляет собой набор визуальных характеристик и какие критерии необходимо использовать при их оценке. Подобие некоторых графических элементов букв шрифта и самих букв шрифта, так же как и шрифта в целом, наводит на мысль о возможности использования для их анализа идей фрактальной геометрии. Согласно Б. Мандельброту: «Фракталом называется множество, для которого размерность Хаусдорфа–Безиковича строго больше его топологической размерности» [9]. В нашем случае, если размерность границы набора букв шрифта представляет собой дробную метрическую размерность, можно говорить, что граница символов шрифта обладает фрактальными характеристиками. Пространственная система должна обладать одинаковыми свойствами во всех масштабах — масштабная инвариантность. Это означает, что при изменении пространственного размера объекта любая количественная характеристика фрактала меняется независимо от этих изменений. Топологическая размерность связана с числом степеней свободы, в то время, как фрактальная размерность интуитивно может быть понята как степень заполнения пространства нерегулярно распределенной субстанцией. При анализе реальных множеств необходимо учитывать, что их фрактальные характеристики могут существовать лишь в ограниченной области масштабов. Соотношение между длиной абриса символа или набора символов шрифта и его площадью может характеризовать показатель шрифта.

В любом семействе плоских фигур, геометрически подобных, но имеющих различные линейные размеры, отношение длины границы фигуры к квадратному корню ее площади представляет собой число, полностью определяемое общей для семейства формой. Частный случай фрактальной размерности d выражается известной формулой, объединяющей

количество объектов n , при помощи которых проводится измерение, и геометрический размер такого объекта a :

$$d = \frac{\log n}{\log \frac{1}{a}} \quad (1)$$

Кроме того, Б. Мандельброт показал, что для фрактальных множеств выполняется равенство, связывающее длину периметра фрактального объекта P и его площадь S :

$$\frac{P^{\frac{1}{d}}}{S^{\frac{1}{2}}} = \text{const} \quad (2)$$

откуда следует, что $S \sim P^{2/d}$.

Рассматривая шрифт, как геометрическое множество, по аналогии с тем, как предложено в [10], можно применить определение компактности этого множества C :

$$C = \frac{P^2}{S} \quad (3)$$

При наличии процедуры измерения периметра и площади набора символов шрифта, возможно рассматривать этот показатель как его пространственную характеристику и/или показатель шрифта. Кроме компактности, удобно использовать коэффициенты кругообразности C_c (4) и изрезанности (нормализованной компактности) C_n (5).

$$C_c = \frac{4\pi S}{P^2} \quad (4)$$

$$C_n = C_c^{-1} = \frac{C}{4\pi} = \frac{P^2}{4\pi S} \quad (5)$$

Векторные графические редакторы, обладающие собственным макроязыком на основе VBA позволяют решить задачу вычисления этих показателей. В настоящей работе использовался макрос *CurveInfo* для пакета *CorelDraw*. Макрос вычисляет периметр (в мм) и площадь (в мм²) векторного объекта. Было сделано предположение, что влияние знаков препинания, специальных символов и т.п. на разборчивость текстовых информации минимально, поэтому для оценки конкретного шрифта необходимо проводить расчеты только над набором букв. В качестве репрезентации шрифта использовался полный набор из 66 заглавных и строчных букв каждого шрифта (см. рис.1). Для получения информации о периметре конкретной буквы необходимо сложить периметр внешнего абриса буквы ($P_{\text{абрис}}$) и (при наличии) периметр внутрибуквенного просвета ($P_{\text{пр}}$). Периметр полного набора из 66 букв (P) равен сумме периметров всех букв набора (6).

Для получения площади конкретной буквы нужно из общей площади, ограниченной внешним абрисом буквы ($S_{\text{абрис}}$), вычесть площадь внутрибуквенного просвета ($S_{\text{пр}}$). Площадь полного набора из 66 букв (S) равна сумме площадей всех букв (7).

$$P = \sum (P_{\text{абрис}} + P_{\text{пр}}), \quad (6)$$

$$S = \sum (S_{\text{абрис}} - S_{\text{пр}}), \quad (7)$$

Был разработан специальный макрос, автоматизирующий процесс разбиения произвольного количества текстовых объектов на набор векторных графических объектов, составляющих абрисы букв и внутрибуквенные просветы.

Процедура измерения

Для проведения измерения был выбран 21 шрифт (прямое светлое начертание) кеглями 12 и 18 пт. Среди них 9 шрифтов без засечек, 11 шрифтов с засечками и один скриптовый шрифт. К каждому шрифту из выбранного набора была применена процедура вынесения всех внутрибуквенных просветов и последующее объединение в отдельные связные векторные объекты всех абрисов заглавных и строчных букв; всех внутрибуквенных просветов заглавных и строчных букв (рис.1). Затем применялся макрос Curveinfo и вычислялись периметр и площадь по формулам (6), (7). Затем по формуле (5) рассчитывалась нормализованная компактность.



Рис. 1. Разъединение набора букв шрифта на абрисы и внутрибуквенные просветы перед расчетом показателя

Можно предположить, что для шрифтов с засечками и без засечек показатель нормализованной компактности будет иметь различный смысл. Поэтому шрифты с засечками и без засечек рассматривались отдельно. По распределению значений нормализованной компактности для групп шрифтов с засечками и без засечек проводилась систематизация: группировка методом экспертной оценки; кластерный анализ с использованием пакета *Statistica* 10.

Результаты и обсуждение

Результаты измерений и расчетов показателей шрифтов по формуле нормализованной компактности (5); для набора шрифтов размером 12 и 18пт. приведены в табл. 1.

Результаты расчета нормализованной компактности для шрифтов кеглем 12 пт., 18 пт.

Таблица 1

№ п.п.	Шрифт	Особенность	C_n , кегль 12 пт.	C_n , кегль 18 пт.
1	Vanta		305.44	305.45
2	Text Book		417.82	417.82
3	Comic Sans		439.25	416.32
4	Verdana		458.56	464.4
5	Avantgard		469.5	469.5
6	Arial		480.52	469.35
7	Century Gothic		574.92	574.92
8	Futuris		605.28	605.27
9	Futura		824.98	824.98
10	Bruskovaya	Засечки	471.46	471.45
11	Antiqua	То же	585.38	585.38
12	Mysl Narrow	—«—	653.17	653.17
13	Times New Roman	—«—	674.68	654.96
14	Baskerville	—«—	681.9	681.9
15	Journal	—«—	702.94	703.73
16	Book Antiqua	—«—	714.04	714.04
17	Baltica	—«—	777.51	792.16
18	Classic Russian	—«—	796.17	796.16
19	Garamond	—«—	875.33	875.32
20	Literaturnaya	—«—	879.76	801.94
21	Art Script	скрипт	1716.72	1650.94

Как можно видеть из табл. 1, нормализованные компактности шрифтов являются практически постоянными величинами, что свидетельствует об объективном характере предложенного масштабно инвариантного показателя, который удобно использовать в исследованиях. Небольшая вариативность предложенного показателя для некоторых шрифтов может быть объяснена ошибками масштабирования.

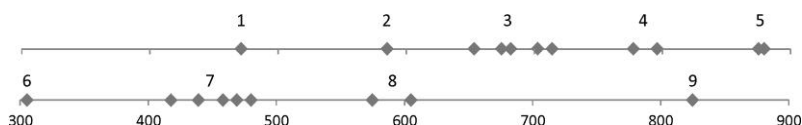


Рис. 2. Распределения значений нормализованной компактности (безразмерная величина, ромбы на графиках) для групп шрифтов с засечками и без засечек. Цифрами 1-9 обозначены группы шрифтов с близкими значениями показателя нормализованной компактности (группировка по экспертным оценкам)

На рис. 2 приведены распределения нормализованной компактности для групп шрифтов с засечками и без засечек (за исключением скриптового шрифта). Группировка проведена при помощи метода экспертной оценки по распределениям, приведенным на рис 2. Скорость чтения шрифтов с засечками (рис. 2) в среднем несколько выше, чем шрифтов без засечек. Скриптовые шрифты, как отмечалось выше, из-за сложности с различением букв обладают низкой скоростью чтения, при этом показатель нормализованной компактности у

них чрезвычайно высок (табл.1, *ArtScript*). В целом предложенный показатель демонстрирует применимость для ранжирования внутри группы однотипных шрифтов.

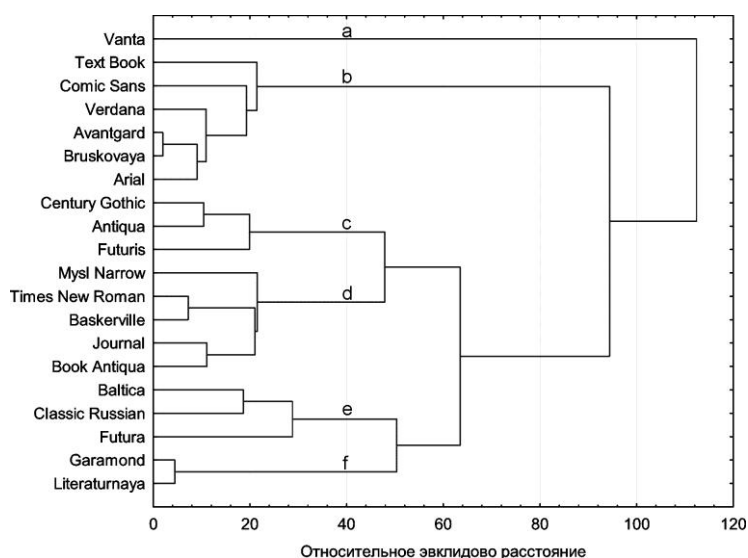


Рис. 3. Кластеризация шрифтов по значению показателя нормализованной компактности

Выводы

Учет предложенного показателя в исследованиях чтения (например, [11]) позволит выявить предикторы скорости чтения и качества усвоения прочитанного не только для бумажных, но и для «электронных» текстовых информационных. Для оценки применимости предложенного показателя необходимы исследования большого количества различных шрифтов и статистический анализ полученных результатов, который может выявить факторы, влияющие на удобочитаемость шрифтов.

Список литературы

1. Ушакова М.Н. Новый шрифт для художественной литературы // Полиграфическое производство. 1959. № 11. — С. 26–27.
2. Александрова Н.А., Чиминова В.Г. Новые шрифты для газет//Полиграфическое производство. 1962. № 5 — С. 57–61.
3. Токарь О.В., Зильберглейт М.А., Литунов С.Н. Оценка удобочитаемости шрифта на материале официального документа//Омский научный вестник. 2009. № 2 (80). — С. 246–249.
4. Тарбеев А.В. Шрифтовое оформления современного учебника// Известия высших учебных заведений. Проблемы полиграфии и издательского дела. 2004. № 4. — С. 93–102.
5. Tinker M.A. Legibility of print. Iowa State University Press, Ames, Iowa. 1963. — P.329.
6. Sanford A.J. The Mind of Man: Models of Human Understanding. New Haven: Yale University Press, 1987. — P.141.
7. McCarthy M.S., Mothersbaugh D.L. Effects of typographic factors in advertising-based persuasion: A general model and initial empirical tests. // Psychology & Marketing, 2002, Vol.19, Issue 7–8. — P.663–691

8. Sanocki T., Dyson M.C. Letter processing and font information during reading: beyond distinctiveness, where vision meets design. //Atten Percept Psychophys. 2012. №74 (1). — P.132–145.
9. Мандельборт Б. Фрактальная геометрия природы. М.: Институт компьютерных исследований, 2002. — С. 656.
10. Алексанина М.Г., Карнацкий А. Ю. Сравнение пространственных характеристик полей морского льда и облачности по данным видимого канала AVHRR/NOAA на примере Охотского моря / Шестая всероссийская открытая ежегодная конференция «Современные проблемы дистанционного зондирования Земли из космоса» М.: ИКИ РАН, 10–14 ноября 2008 г. Сборник тезисов конференции. — С. 299–302.
11. Тарасов Д.А., Сергеев А.П., Корнилова Ю.И. Интерлиньяж как фактор скорости чтения на примере бумажных и веб-текстов// Известия высших учебных заведений. Проблемы полиграфии и издательского дела. 2013. № 2. — С. 81–88.